NFL Adjusted Plus-Minus Special Teams Models

Metrics Assessing Player's EPA and Penalty Contribution During Punts

Consultant: Matt Johnson
Client: NFL Big Data Bowl
Advisors: Dr. Steve Marron, Ph.D. & Dr. Perry Haaland, Ph.D.

26 March, 2022

1. Abstract

2. Introduction

This paper introduces two new metrics for evaluating player performance during punting situations in the NFL: Regularized Adjusted Expected Points Added Contribution (RAEPAC) and Regularized Adjusted Penalty Contribution (RAPC). To build these new metrics, I adapt a statistic historically used in the NBA called Adjusted Plus-Minus (APM). This metric adjusts the conventional Plus-Minus metric to control for every player on the court during a stint where the same 10 players are on the court. In the NFL punt return case, this is simply the 22 players on the field during a punt. In classic APM for basketball, the variable of interest is the points scored during the stint, however this will not work for football since points are rarely scored during punts. In RAEPAC, I use Expected Points Added (EPA) as the variable of interest, while in RAPC I use a penalty indicator coded as a `-1` if the punt team commits the penalty, a `1` if the return team commits the penalty, and a `0` if there is no penalty. EPA calculation is further explained in the Methodology section. Specifically, to calculate RAEPAC I use Ridge regression to to get regression coefficients for each player that was on the field during a punt. To calculate RAPC, I use a Regularized Ordinal Logistic Regression to get coefficients for each player. The regression coefficients for the respective models are what define these two new metrics. The model choices will be discussed further in the Methodology section. These metrics can be used to rank players regardless of their position on how much they positively contribute towards the success of their team on punting situations. This type of evaluation method could be extremely useful for on-field decision makers when deciding who should be on the field during crucial punting plays. Also, these metrics can be aggregated on a team level to rank how successful teams are during punts. Assessing players regardless of their position could change the way coaches think about their personnel and shed light on players that may not get enough credit for their contribution.

2.1 Data

The final dataframe used for analysis has 5890 plays with 2005 unique players and two response variables: EPA and the penalty indicator. Each row of is a play from 2018-2020 where the team lined up to punt the ball away and each column is a player that could have been on the field. Each cell of the dataframe represents if the player was on the field during that play or not. A player on the punting team is given a `-1` and each player on the receiving team is given a `1`. Players that were not on the field during that play are coded as a `0`. This follows the methodology first employed by Dan Rosenbaum for computing APM is basketball (Rosenbaum).

This dataframe is inherently rank deficient and therefore I cannot use standard linear regression, rather I use a penalized approach so the sample covariance matrix is invertible. The dataframe is rank deficient because if I remove a players column, I still know whether he was on the field on offense or defense by simply using the rest of the columns. According to this, the matrix does not have full rank. In other words, the rank of the matrix will always be at most N-1 given there are N players represented. Below is a glimpse of the dataframe.

PlayId	EPA	Pen.Yrds	32361	38707	40008	40191	40408	42450	42515
2018090600_366	0.75865422	-1	-1	1	1	1	-1	1	-1
2018090600_872	-0.41094785	0	0	-1	0	-1	1	0	1
2018090600_973	-0.51340565	0	-1	1	0	1	-1	1	-1
2018090600_1850	-0.19499974	0	0	-1	0	-1	1	0	1
2018090600_1989	-0.34500556	1	0	-1	0	-1	1	0	1
2018090600_2406	0.08443995	0	0	-1	0	-1	1	0	1

3. Methodology

In this section I discuss the EPA variable calculation, Exploratory Data Analysis (EDA), and model selection process. I use a similar methodology as explained in Dan Rosenbaum's 2004 paper, titled "Measuring How NBA Players Help Their Teams Win", where he describes the creation of the APM model (Rosenbaum).

3.1 EPA Calculation

EPA is calculated by taking the difference between the expected points before and after the play occurs. Using historical data, the amount of points a team is expected to score on the current drive depending on the down, distance to the first down, and field position can be calculated. Expected points is based off the notion that not all yards gained are the same. For example, a 2-yard gain on 4th down and 3 does not increase a teams chance of scoring, while a 2-yard gain on 4th and 1 does because the offense will remain on the field. Expected points quantifies this difference by relating each play to how much it changes the chances of scoring on that drive. Figure 1 shows how expected points changes as a team drives down the field towards the opponents end zone (Greer).
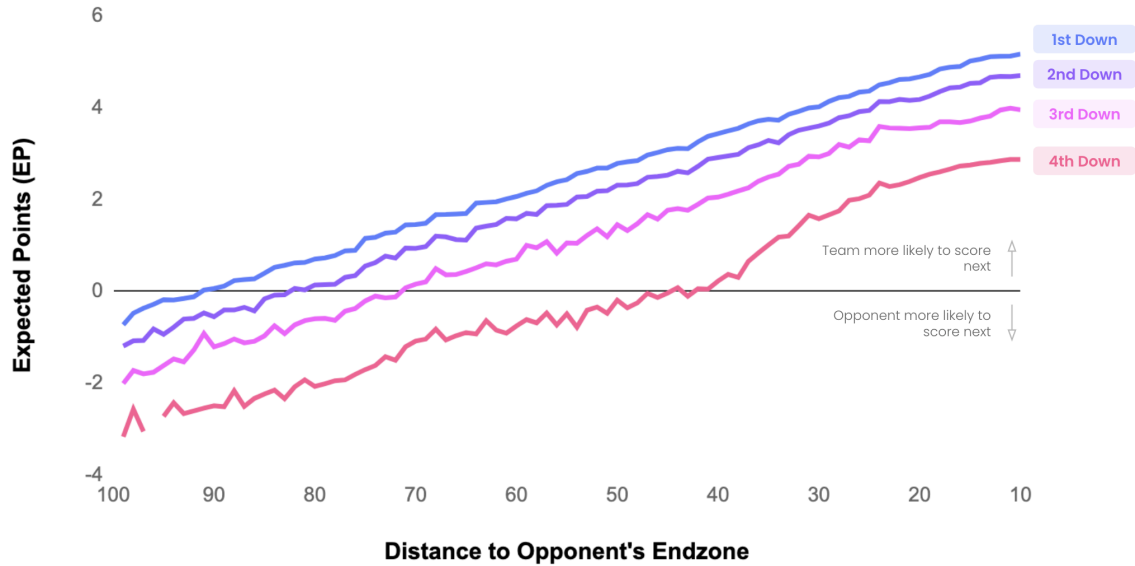
Figure 1: This plot shows how expected points changes depending on the down and distance to the end zone. As shown, it is far better to be closer to the end zone on early downs.

As can be seen, expected points increases as the team moves closer and closer to the end zone. Also, the down has a significant impact on the expected points as 1st down will always be higher than 4th. When expected points is negative, this means the opponent is more likely to score on the next drive because the team currently in possession is likely to punt. This situation occurs when a team is near their own end zone. EPA shows how much the expected points changed because of the play that happened. When EPA is positive the play is deemed successful for the offense. I'll use EPA, calculated by the nflfastR play-by-play database, as my main response variable because it directly reflects the success of the play in terms of expected points scored in the next drive. The punting team will want this variable to be as negative as possible, while the return team will try to increase it to nearly 7.

3.2 Exploratory Data Analysis

I complete a preliminary EDA to validate and visualize the data before the modeling procedures. Completing and EDA prior to modeling is a crucial step that ensures the data are correct and the model choices are valid. I begin by making certain the final dataframe is correct. To complete this I first check that each row sums to zero to show there are the same amount of offensive and defensive players on the field. I then check to see that there are always 22 players on the field. Finally, I choose 500 random plays and check to make sure the correct players are on the field.

Next, I visualize the response variables: EPA and penalties. This step helps to decide what kind of models

are appropriate for these responses. Below I plot the histograms of the two variables.
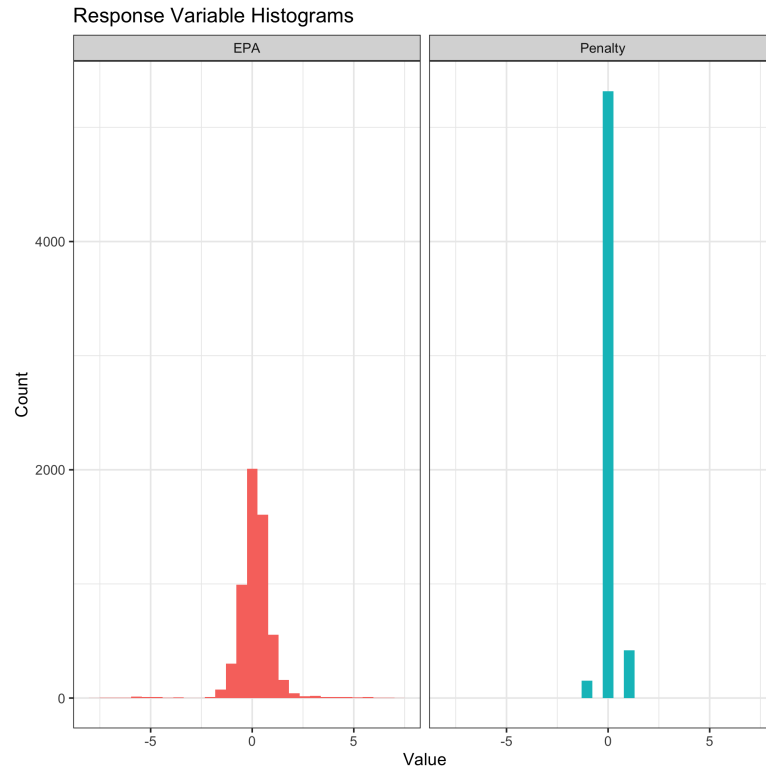
Figure 2: This plot shows the distribution of the two response variables. It seems the EPA response is approximately normal around zero. On the other hand there are a lot of plays with no penalties.

As seen in Figure 2, there seems to be an inflated amount of zeros in both distributions. Penalties is a categorical variable and there are many plays that do not have any penalties and thus have a value of zero. On the other hand, EPA is a continuous variable and many of the values are close to zero, indicating that the punt leads to neither team gaining an advantage. Punts that result in a fumble and the punting team recovering the ball result in a very negative EPA meaning the punting team was extremely successful and is now likely to score. Similarly, punts that result in a huge return have a high EPA close to 7 that shows the return team was successful. In Figure 3, I show box-plots of the response variables to reinforce these findings.
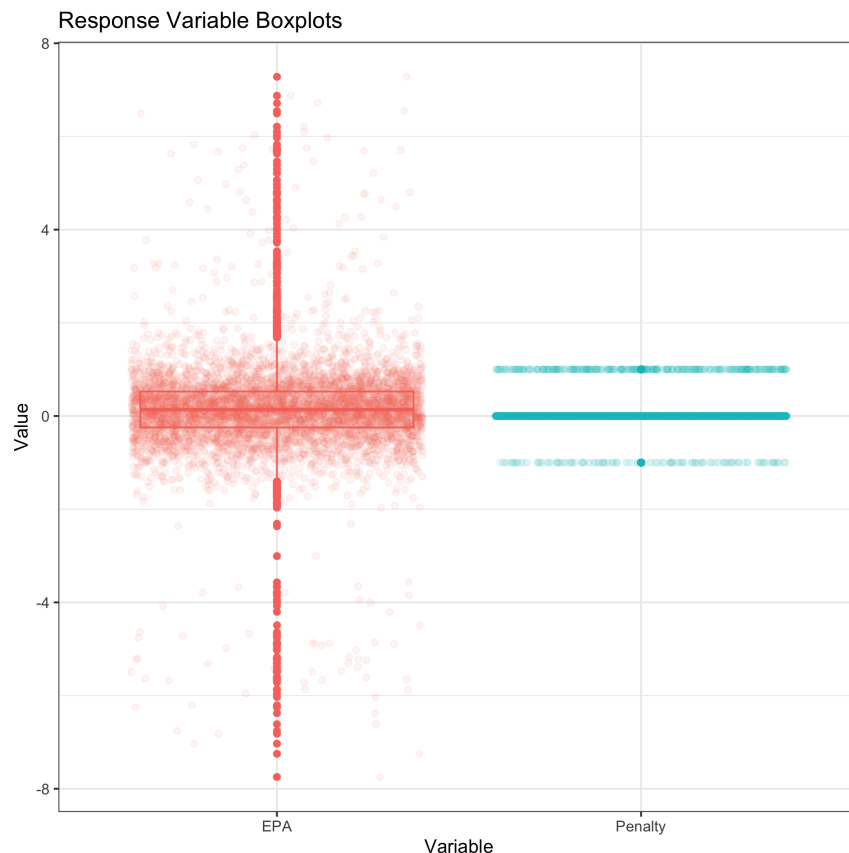
Figure 3: These box-plots help reinfroce what is seen in Figure 2. The mean for both response variables is right around zero and EPA seems approximately normal.

Figure 3 strengthens the observation that EPA is approximately normally distributed around zero and has small tails at 7 and -7. Penalties only takes three categorical values that correspond to the type of penalty committed. A vast majority of plays do not have a penalty and thus have a zero value for this response.

The last visualization I make for the EDA shows an rudimentary look at the impact players have while they are on the field. Figure 4 shows the average EPA and penalty yards for each player when they are on the field against the population average for both response variables. This plot helps to explain the intuition behind this analysis, I suspect that players with higher averages also may contribute more.
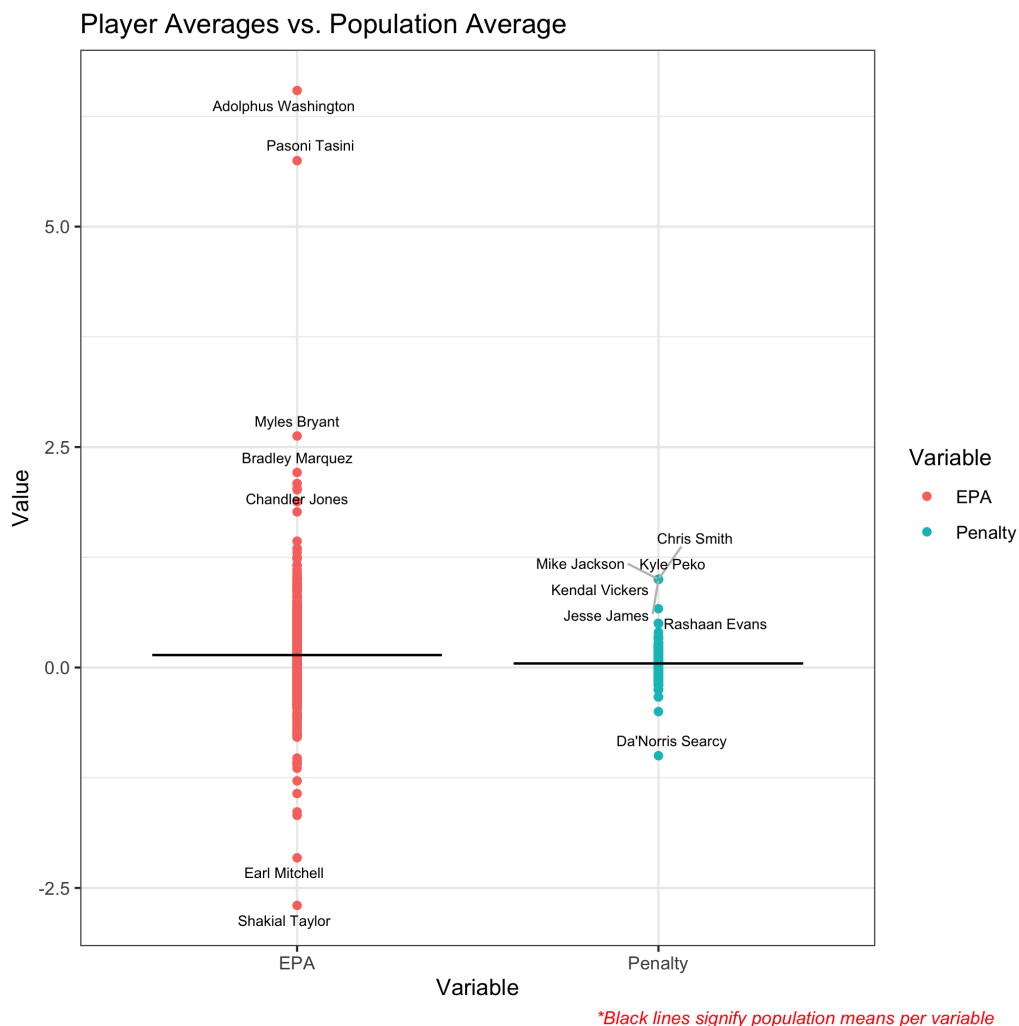
Figure 4: This plot shows how individual players averages are different than the population average for both response variables. Lablued players are 2 standard deviations away from the population mean.

In Figure 4, I find players that have skewed averages because they were on the field for very few plays. Adolphus Washington only played one snap on the return team that happened to be a major return, thus resulting in a high EPA. Similarly, Da'Norris Searcy and Kendal Vickers each played a single snap that happened to have a penalty. Because of this, these players with low snap counts will not be considered during the contribution analysis as they have not played enough snaps to truly assess their contribution. During the modeling procedures I set the minimum snap count to 25 meaning a player will need approximately three games on the punt or return team for their contribution to be considered noteworthy.

3.3 Model Selection

The model selection process follows directly from the results for the EDA. I attempt to use models that are

appropriate for the distribution and real world interpretation of the two response variables. In each case, the regression coefficient for each player will represent their individual contribution controlling for all other plays on the field during that play.

EPA Model: As discussed in the Data section, the player interaction matrix is rank deficient and thus I must use a penalized approach to the regression. Adding the penalty term ensures the player interaction matrix is invertible. EPA is approximately normal around zero and continuous, so it is reasonable to use Ridge regression to account for the rank deficiency and still find reasonable coefficients for each player to represent their contribution. I decide to use Ridge here rather than LASSO because I do not want any player to end up with zero as their contribution. If I had used LASSO, many of the player's regression coefficient would likely be shrunk to zero and I would not be able to rank them. This penalized approach also controls for any multicollinearity that may in the player matrix. Many players, such as punters and long snappers, are likely always on the field for the same plays which may introduce multicollinearity that would inflate coefficient values. I use 10-fold cross validation to find the optimal penalty parameter. The best model from these cross validation procedures is used to find the EPA contribution for each player.

Penalty Model: This response is a three factor ordinal categorical variable. Like in the EPA model formulation, I need to use a penalized approach since the player interaction matrix is rank deficient. Thus, I decide to use a Regularized Ordinal Logistic Regression. This regression technique works as a classification model when the response has three or more categories where the order matters and adds a penalty parameter to ensure the player interaction matrix is invertible. In this model I use an LASSO penalty because I do not need to rank those players that do not significantly contribute towards penalties. In the EPA model I assessed how each player contributes to EPA since every play has a measured EPA, but in this case I only care about those players that actually contribute to the penalties. With a LASSO penalty, players on the field during plays that did not have a penalty will have a contribution of 0. I use 10-fold cross validation here to find the optimal penalty parameter. The model with the highest out of sample classification accuracy is used to find each player's penalty contribution.

4. Results & Discussion

In this section I discuss the results from the two regression analyses. I present the result for the player level as well as the aggregated team level to assess both individual player and team effectiveness during punt plays.

4.1 Player Level Analysis

First, I'll discuss results on the individual player level for both RAEPAC and RAPC. Here I assess how each

player contributed towards EPA and penalties controlling for all other players that were on the field with them during a stint. Figure 5 shows the top 10 and bottom 10 players for the RAEPAC metric.

Top 10 & Bottom 10 Player EPA Contribution

2018-2020 | Minimum of 25 punt returns in that time
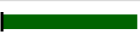
PLAYER			SNAP COUNT	EPA CONTRIBUTION
John Cominsky		DT	36	0.0198
Collin Johnson		WR	26	0.0172
Jordan Miller		CB	25	0.0149
Jason McCourty		CB	53	0.0140
Bruce Miller		FB	29	0.0138
J.T. Hassell		SS	27	0.0135
Paul Worrilow		OLB	33	0.0129
Darius Slay		CB	37	0.0128
Ian Thomas		TE	44	0.0120
Peyton Barber		RB	63	0.0119
Wayne Gallman		RB	25	-0.0117
Chase Allen		OLB	35	-0.0118
Martrell Spaight		ILB	36	-0.0136
Van Jefferson		WR	38	-0.0141
Albert Wilson		WR	25	-0.0173
Ryan Kerrigan		DE	27	-0.0189
Da'Ron Payne		DT	27	-0.0195
Tyron Johnson		WR	28	-0.0196
Kerry Hyder		DE	26	-0.0199
Arik Armstead		DE	32	-0.0218

Figure 5: This figure shows the top 10 and bottom 10 players with regard to EPA contribution. The bottom 10 players are those with negative contribution with Arik Armstead being the worst.

The top 10 players contribute positively towards EPA meaning they contribute efficiently and help their team be successful. On the other hand, players in the bottom 10 contribute negatively meaning their opponent is more successful during the punt return. Players in the bottom 10 effectively detract from their team's success. Because of the regularization technique, these contribution metrics cannot be interpreted in the same way as in ordinary least squares regression coefficients. Instead, they must be interpreted as the amount of differential contribution with respect to the penalty parameter. In other words, the contribution is an approximation scaled by the penalty parameter.

It is interesting to see that no position is specifically better than any other. Usually, the ball carrier gets

most of the credit during a punt return, but this metric does not discriminate against position and attempts to highlight those players that normally get little to no credit during returns. Also, snap count does not seem to have a major impact on contribution. Players that are on the field more (or less) do not seem to have better (or worse) contribution.

Figure 6 shows a similar plot for the RAPC metric. In this case, I only plot the top 20 players because decision makers will want to know the players that contribute most towards penalties. For this metric, positive contribution means the player contributes towards penalties that hurt their own team. In other words, positive contribution is bad unlike in RAEPAC where positive contribution means they contribute towards success.

Top 20 Player Penalty Contribution
2018-2020 | Minimum of 25 punt returns in that time

PLAYER			SNAP COUNT	PENALTY CONTRIBUTION
Brian Allen		CB	44	1.7943
Dekoda Watson		OLB	26	0.9465
Donnie Jones		P	44	0.6681
Josh Hawkins		CB	28	0.5777
Michael Davis		CB	96	0.5690
K'Waun Williams		CB	25	0.4456
James Proche		WR	47	0.4154
Chapelle Russell		LB	37	0.3997
Trevor Daniel		P	85	0.3971
Isaiah Johnson		S	86	0.2764
Edmond Robinson		OLB	48	0.2699
Brandon Graham		DE	25	0.2616
Javelin Guidry		DB	36	0.2516
D'Juan Hines		LB	100	0.2468
DeShawn Shead		CB	77	0.2394
Ray-Ray McCloud		WR	111	0.2380
Coty Sensabaugh		CB	42	0.2168
Jack Cichy		ILB	88	0.2004
Patrick Chung		SS	27	0.1833
Deon Bush		FS	208	0.1775

Figure 6: This figure shows the top 20 players in penalty contribution. These players detract from team success as positive contribution means the penalties hurt their team.

Like in EPA contribution, Figure 6 shows that player position does not seem to be a contributing factor. This

makes perfect sense as any position can cause penalties. Though, it is interesting to see that there are a few punters here. I assume these players are not the ones causing the penalties as punters rarely are penalized. Rather, they are simply on the field when penalties happen. Similarly, there are a few players (James Proche & Ray-Ray McCloud) that are the return men for their respective teams. It is usually unlikely that the ball carrier will be penalized so they are likely just on the field when their team is getting penalized, thus the penalty contribution is attributed to them. This discovery points to a flaw in the player level penalty analysis. Players that are likely not getting penalized are being shown as high penalty contributors because they are on the field the most during stints when penalties occur. Rather than attributing these penalties to specific players, I'll move on the aggregate on a team level to assess which teams have high penalty contribution. Players in Figure 6 may not truly be contributing to penalties, rather this plot points to teams that may have high penalty contribution.

4.2 Team Level Analysis

5. Conclusion & Future Research

5.1 Conclusion

5.2 Future Research

6. References

1. Greer, R. (2021, October 23). What are expected points added (EPA) in the NFL. nflelo. Retrieved March 21, 2022, from https://www.nfeloapp.com/analysis/expected-points-added-epa-nfl

2. Rosenbaum, D. (2004, April 30). Measuring How NBA Players Help Their Teams Win. Picking the difference makers for the All-NBA Teams. Retrieved March 21, 2022, from http://www.82games.com/comm30.htm

3.

Appendix

A full appendix can be found on github here: https://github.com/mattymo18/2022-NFL-Big-Data-Bowl